CAAI Applied Machine Learning Pre-Doc Program
Analysis Task

Instructions:

Congratulations on making it to the second phase of the application process! While we have screened CVs and application forms, your submission to this task will form the basis of our decision to move forward with your application. Please spend time on this task to showcase your approach to problem-solving, your ability to communicate, and your work ethic as we can expect to see it in action.

Due Date: November 20, 2020

Submission link: https://forms.gle/tagRjeGCQBHVZNiR9

What we are looking for:

- An analytical approach to problems
- Innovative solutions for real-world contexts
- Strong communication skills
- Deep technical knowledge of appropriate models and tools

The work you will be doing in this role is varied, but the common element will be its possible impact in real world situations. The ability to understand the problem, the solution, and the context of each is hugely important.

Tasks:

1. Find the The IMDB-WIKI dataset
2. Plot the age distribution of this population, and determine the size of the bucket “15 to 25 years old”
3. What percentage of this population is "30 year old males"?
4. Using this dataset, how would you build an algorithm which can predict a person's age using their face?
5. What are the tradeoffs of the algorithm you selected? What is the loss function?
6. Would you trust this algorithm in high stake situations (for instance to detect minors at the airport)? Why/Why not?
7. Write the algorithm flowchart, and provide a legend explaining each step in the flowchart. How would you train this algorithm? How would you test the algorithm?
8. Summarize all your work from part 1-7 in a document. Include as much detail as possible, and Github link for any written code